What Works Clearinghouse

IES Institute of Education Sciences

Topic report Curriculum-based interventions for increasing K-12 math achievement—middle school

Updated December 1, 2004

Purpose

This report reviews the available evidence from research conducted since 1983 on the effectiveness of curriculum-based interventions for improving mathematics achievement for middle school students.

Evidence base

From a systematic search of published and unpublished research the WWC identified 10 studies of the effects of 5 middle school math interventions that met WWC standards of evidence.

Students and interventions

Most students were in the 8th grade, with a range from 6th to 10th. Each intervention was the main curriculum in the study reviewed.

Outcomes

Most studies reported findings for one or more standardized mathematics tests.

Effectiveness based on randomized controlled trials

Two of the four randomized controlled trials showed evidence of having significantly increased achievement—*Cognitive Tutor®* and the *I CAN Learn® Mathematics Curriculum.*

Effectiveness based on quasi-experimental design studies

No quasi-experimental design studies had statistically significant effects.

This topic report covers only interventions with studies meeting WWC evidence standards or meeting standards with reservations. We acknowledge that schools choose a curriculum after considering many factors. We believe that it is useful to identify what studies of achievement show, even if for a small subset of all studies.

Executive summary

Evidence base

From a systematic search of published and unpublished research, the What Works Clearinghouse (WWC) identified 10 studies of 5 curriculum-based interventions for improving mathematics achievement for middle school students. These include all studies conducted in the past 20 years that met WWC standards for evidence.

The five curricula having at least one study of effectiveness that meets WWC standards for evidence (see below) include *Cognitive Tutor®, Connected Mathematics Project, The Expert Mathematician, I CAN Learn® Mathematics Curriculum,* and *Saxon Math*.

The WWC identified 66 other studies that included evaluations of 15 additional interventions. Because none meets the WWC standards for evidence, we cannot draw any conclusions about the effectiveness of these other 15 interventions. The WWC also identified an additional 24 interventions that did not appear to have any evaluations.

WWC standards for evidence

WWC categorizes studies by the causal validity of the findings. Randomized controlled trials (RCTs) without serious implementation flaws meet the WWC evidence standards for causal validity. Such studies ensure that there are no systematic differences between intervention and comparison groups. Quasi-experimental design (QED) studies that include pretest measures of the outcome of interest and/or that match the intervention and comparison groups on important characteristics and RCTs that have notable implementation flaws (such as high sample attrition) meet the WWC evidence standards with reservations. For both there is greater uncertainty about the causal validity of the effectiveness estimates.

The other major dimension of evidence standards pertains to generalizability. Interventions evaluated on the basis of a scant number of small studies have limited generalizability; those based on one or more large studies, or on many smaller studies, have greater generalizability.

Quality of the evidence base

Only 5 of the more than 40 middle school math interventions known to be available for adoption have any studies of their effectiveness that meet the WWC evidence standards.

Only four RCTs meet the WWC standards for causal evidence without reservation—one each for *Cognitive Tutor®, The Expert Mathematician, I CAN Learn® Mathematics Curriculum,* and *Saxon Math.* Each focused on a different intervention, and two studies (*The Expert Mathematician* and *Saxon Math*) use very small samples. *I CAN Learn® Mathematics Curriculum* and *Saxon Math* have also been evaluated using QED studies, as has the *Connected Mathematics Project.*

The overall findings are summarized in table 1.

Cognitive Tutor®: Only one study of this intervention meets the WWC standards for evidence, an RCT that included 360 9th-grade regular education students (in the subsample that was randomly assigned and took the Educational Testing Service [ETS] test). Students using *Cognitive Tutor®* were compared to students using *McDougal Littell's Heath Algebra I.* The intervention group scored, on average, 0.23 standard deviations higher than the comparison group on the ETS Algebra test, a statistically significant difference. In other words, the average intervention group student scored at the 59th percentile relative to the comparison group when the average control group score was set at the 50th percentile.

Connected Mathematics Project: Three QED studies of this multiyear curriculum meet the WWC standards for evidence with reservations. The results of the three studies are inconclusive. Two estimated sizable differences in mean scores in favor of the intervention group (0.32 and 0.43 standard deviations). But the statistical significance of the estimate based on the large sample of students could not be determined since the analysis is at the wrong level and estimates from the smaller study are not statistically significant. The third study found a negative effect, which was not statistically significant.

The Expert Mathematician: The only study of *The Expert Mathematician* is a small RCT focused on 8th grade regular

Executive summary *(continued)*

education students (70 students). It compared math achievement for students using *Transition Mathematics*, part of the University of Chicago School Mathematics Project. The study found that *The Expert Mathematician* outperformed *Transition Mathematics;* however, statistical significance could not be calculated from the data in the study.

I CAN Learn® Mathematics Curriculum: Three studies of this intervention—one RCT (254 students) and two QED studies-meet WWC evidence standards or meet evidence standards with reservations. All include 8th-grade students, and some also include other grade levels. All resulted in some estimates of math test effects that are positive for the intervention group, although only one was significant. The RCT suggests that the intervention had a significant positive effect on math test scores, a difference of 0.41 (in standard deviations) in favor of the intervention group. In the RCT, the average intervention group student scored at the 66th percentile relative to the comparison group when the average comparison group score was set at the 50th percentile. One QED study found that *I CAN Learn®* students scored higher than comparison students, but the effects were not statistically significant. A second QED study analyzed the data at the wrong level (student, not classroom), making it impossible to determine statistical significance.

Saxon Math: Two small studies, both focusing on 8th-grade students, meet WWC evidence standards or meet WWC evidence standards with reservations—one RCT and one QED study. In both studies the students using *Saxon Math* scored higher (but not significantly higher) on math achievement tests than did students using another curriculum—in one case the *University of Chicago Mathematics Project* curriculum and in the other a Scott-Foresman text for most of the comparison sample. In the RCT, the difference in raw test scores is 0.11 standard deviations in favor of the intervention group. That is, the average intervention group student performed at the 55th percentile relative to the comparison group when the average comparison group score was set at the 50th percentile. This difference is not statistically significant. The difference is 0.41 in the QED study with a sample of 78 students. This QED study analyzed the data at the wrong level, making it impossible to determine the statistical significance of the estimated effect.

Improving the evidence

The evidence base is sparse. There have been few randomized controlled trials of math interventions for middle school students, and those few trials conducted tend to be small.

Table 1 Summary of evidence of effectiveness of middle school mathematics curricula

Intervention	*Study sample and setting*[a]	*Strength of the evidence*		
		Strength of studies	*Number, size of studies*[b]	*Estimated effects*[c]
Cognitive Tutor®	9th-grade regular education students	RCT, meets standards	1 RCT • 360 students[d]	RCT • Percentile rank: 9.1 • Standard scores: 0.23
Connected Mathematics Project	6th–8th grades; Midwest, west, east, northeast, south regions; rural, urban, suburban schools; both low and high socio-economic areas	QED, meets standards with reservations	3 QEDs • 1,095 students • 50 schools • 36–42 schools[e]	QED • No significant effects[f]
The Expert Mathematician	8th-grade, low-income students	RCT, meets standards	1 RCT • 70 students	RCT • No significant effects[g]
I CAN Learn® Mathematics Curriculum	7th–10th grades, urban Title I schools	RCT, meets standards QED, meets standards with reservations	1 RCT • 254 students 2 QEDs • 169 classes • 116 classes	RCT • Percentile rank: 15.9 • Standard scores: 0.41 QED • No significant effects[f]
Saxon Math	8th-grade; urban school	RCT, meets standards QED, meets standards with reservations	1 RCT • 36 students 1 QED • 78 students	RCT • No significant effects QED • No significant effects[f]

a Some study authors did not provide full information on some aspects of the sample and setting. Samples and settings for the studies may be broader than reported here.
b Size of entire sample for each study reported in the units of assignment; note that some analyses, including those profiled in this report, may use smaller subsamples.
c Only statistically significant effects reported here. Gains relative to comparison group.
d Subsample size for students who were randomly assigned for analysis of effects measured with ETS Algebra test.
e Each of 3 cohorts had a different number of schools.
f It was not possible to accurately compute significance levels in one study, as the unit of analysis did not match the unit of assignment.
g It was not possible to accurately compute significance levels.

This report

After a thorough and critical review of the research, the WWC identified five middle school math curricula that have at least one evaluation meeting WWC evidence standards (see table 1).

Two of the four RCTs examining the effects of middle school mathematics curricula show statistically significant, favorable effects for the intervention group relative to their comparison group, who were taught following the regular school math curricula. The estimated effect of *Cognitive Tutor®* in the one RCT is roughly equivalent to the average intervention group student gaining about nine points in the percentile rankings relative to the average comparison group student—a difference denoted by the end of the wide bar in figure 3. Given the study design and this estimate of the effect, we are 95% confident that the intervention group earned test scores that, on average, placed them between 1 and 17 percentile points above the average student in the comparison group, as indicated by the black line through the wide bar denoting the estimate. For *I CAN Learn® Mathematics Curriculum,* the one RCT estimated that the average test scores for the intervention group increased by an amount that is roughly equivalent to gaining 16 percentile points in the percentile rankings relative to the average comparison group score. We have 95% confidence that the true difference in test score gains for those in the intervention group fall between 6 and 25 percentile points. *The Expert Mathematician* increased achievement, compared to the students using *Transition Mathematics*, but it was not possible to compute statistical significance of this difference. There is no evidence from the one RCT of *Saxon Math* that it had an effect on performance.

This report describes the background for this review, the evidence base, the samples, the outcome measures, and the findings.

Background

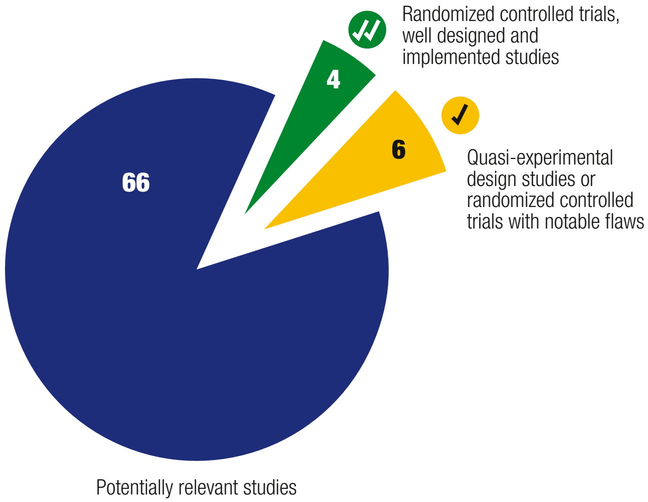
Over the last two decades educators and policymakers in the United States have registered concerns about the need to increase the proficiency of students in mathematics. These concerns have been stimulated, in part, by results of the Second and the Third International Mathematics and Science Studies (SIMSS and TIMSS), reports of an undersupply of mathematicians and scientists (National Research Council 1990), and the importance of mathematics education for the economic health of the nation (National Commission on Excellence in Education 1983). The call to improve mathematics education has also been driven by the widespread belief that competence in mathematics enables individuals to become informed and competent in a technology-dependent society. But that call for improvement has not been accompanied by evidence-based recommendations for how to achieve it.

This WWC Topic Report summarizes evidence from studies that estimate the effects of interventions for improving the mathematics proficiency of middle school students—and that meet WWC evidence standards, usually with some reservations. It covers only interventions that entail materials-based curricula for students in grades 6–9 or any curricula labeled as appropriate for middle school. The interventions focus on one or more of five content areas: number and operations, algebra, geometry, measurement, and data analysis and probability. And they also explicitly promote skills in one or more of the following five areas: problem-solving, reasoning and proof, making connections, oral and written communication, and uses of mathematical representations.

The review addresses two main questions:

1. Are there curriculum-based interventions for which there is evidence of increasing middle school students' learning of mathematics content and skills (that is, what students should know and be able to do? If so, what are they)?
2. Is there evidence that interventions are more effective for certain subsets of students than for others—for example, for students who lag behind in mathematics achievement or those at particular grade levels?

Figure 1 Evidence base for curriculum-based interventions to increase mathematics achievement among middle-school students



WWC standards of evidence

Studies must be designed to permit valid causal inferences about the effects of the intervention. The focus is on randomized controlled trials (RCTs), the only study design that, when well implemented, ensures that there are no systematic differences between intervention and comparison groups at the outset.

WWC also includes, with reservations, studies that use quasi-experimental design (QED) studies and have a reasonable degree of matching of the intervention and comparison groups and that control statistically for pre-intervention differences in the relevant characteristics of the two groups.

Five other features of design and implementation are considered in reviewing studies, but they generally do not result in a decision to exclude a study from the reported evidence base. These are: fidelity of the intervention implementation, relevance of the outcome measures used, generalizability of the findings, strength of the analytical methods used, and the adequacy of the reporting on the study and its results. (www.whatworks.ed.gov/reviewprocess/study_standards_final.pdf).

Evidence base

The WWC identified 44 curriculum-based interventions for middle school students (see appendix table A1) through various sources. A systematic search for studies on the effectiveness of middle school math curricula identified 76[a] studies that warranted further screening (figure 1). Of these, 10 studies of five interventions were judged to meet or exceed the WWC's minimum standard for credible causal evidence of effectiveness.

The WWC identified no studies that meet standards for the other 39 curriculum-based interventions. Outcome studies for 15 of these interventions were identified, but none of them meets WWC evidence standards. For the 24 other interventions, no outcome studies were identified. This lack of evidence does not mean that the interventions are ineffective—just that they are not yet proven.

WWC groups studies into two levels of evidence: randomized controlled trials (RCTs) and quasi-experimental design (QED) studies. The five interventions evaluated through one or more of the studies in this review are *Cognitive Tutor®* (one RCT), *Connected Mathematics Project* (three QEDs), *The Expert Mathematician* (one RCT), *I CAN Learn® Mathematics Curriculum* (one RCT and two QEDs), and *Saxon Math* (one RCT and one QED) (figure 2).

The 10 studies in this review differ in their study design and other considerations—and, indeed, all have some shortcomings. Only four of the interventions have been tested in an RCT, and none has been tested in more than one RCT. Studies were relatively small in scale, and several analyzed data at the student level even though classrooms were the unit of assignment to intervention.

Three of the four RCTs and only one of the six QED studies documented that the interventions were reasonably well implemented. None of the studies provided estimates of intervention effects for key subgroups of students, such as those defined by gender, race, or socioeconomic background. Some of the studies report findings separately for students in different grades or enrolled in different math classes. There is some diversity in the community context, but the evidence base is not large enough to permit meaningful comparison of findings by school or community context.

Most studies reported only a simple difference between means. QED studies that used matching or statistical controls to adjust for pre-intervention characteristics often failed to provide full reporting of the multivariate results. Incomplete reporting makes it difficult to fully assess the studies. All 10 studies reported estimates of intervention effects on relevant outcomes based on standardized tests or other measures with face validity. Studies that did not do so were screened out.

a The research findings from some studies have been reported in more than one report or article. In this review, we report all the findings from a single study together, regardless of the citation for which the findings were presented.

Figure 2 Studies of interventions to increase mathematics achievement among middle-school students, by study design and implementation

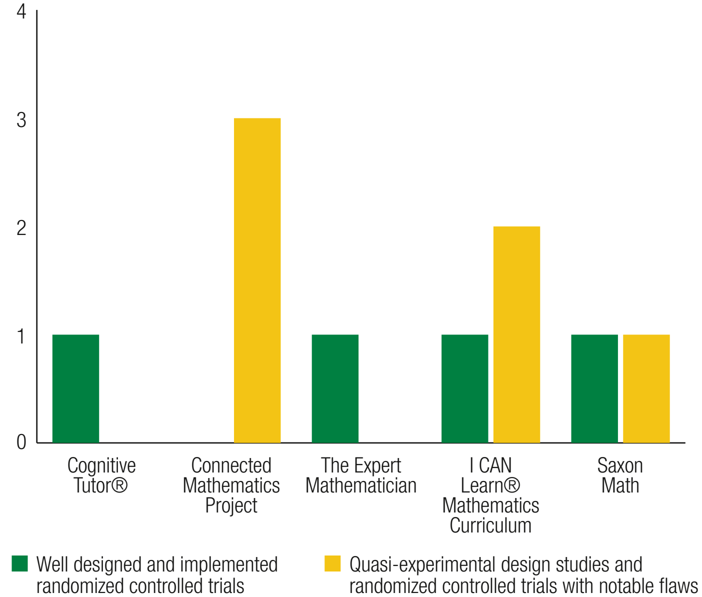


Table 2 Characteristics of the interventions

Intervention	*Grades targeted*	*Intensity*	*Software dependent*
Cognitive Tutor®	9th	Daily for 1 year	Yes
Connected Mathematics Project	6th–8th	8 units a year for up to 3 years	No
The Expert Mathematician	8th	170 minutes a week for a year	Yes
I CAN Learn® Mathematics Curriculum	7th–10th	109 lessons a year	Yes
Saxon Math	8th	120 hours a year	No

Students and interventions

The studies covered here focused on interventions designed to increase mathematics achievement among middle school students who did not have any identified special education need. Most frequently, the students in the intervention groups were in 8th grade (table 2). But, the study of *Cognitive Tutor®* focused on 9th graders. Two of the three studies of *Connected Mathematics Project* focused on 6th through 8th graders. And one of the studies of *I CAN Learn® Mathematics Curriculum* focused on 7th through 10th graders.

One of the RCTs (*Saxon Math*) focused explicitly on "mathematically talented" students. Two of the QED studies (*Connected Mathematics Project* and *I CAN Learn® Mathematics Curriculum*) separated their results by the level of mathematics course students were enrolled in.

The interventions varied in intensity, duration, and how much they relied on technology. To be eligible for this review, the intervention had to be a substitute for the regular curriculum rather than a supplement. Each curriculum considered here was used as the main curriculum in one or more of the studies reviewed. All but *Connected Mathematics Project* and *Saxon Math* relied on technology as an integral part of the intervention. Further details of the interventions and study characteristics are presented in table A2.

Outcomes

Most of the studies reported findings for one or more standardized mathematics tests whose reliability and statistical properties are documented by the test developer. Some studies also reported other outcome measures, such as math grades at the semester's end or performance on locally developed tests, which offer valuable information but often are difficult to interpret. For this reason, estimates of intervention effects on standardized math tests are featured, when they are available. (See Intervention Reports for findings for all relevant outcome measures reported in the 10 studies.)

This Topic Report discusses the estimated effects of the various interventions based on RCTs separately from estimates based on QED studies or RCTs with notable flaws, such as high sample attrition. RCTs, when properly executed and analyzed, offer the strongest evidence about whether the

Outcomes *(continued)*

interventions caused observed differences in the outcomes of interest. Furthermore, current methodological research shows that QED studies do not reliably mirror the results obtained from RCTs for the same intervention and study sample.

Estimated effects: randomized controlled trials

Two of the four RCTs showed evidence of having significantly increased the mathematics achievement of students beyond that of the instruction in the control group—*Cognitive Tutor®* and *I CAN Learn® Mathematics Curriculum* (table 3).

On average, students in the *Cognitive Tutor®* intervention group (randomized sample) scored 0.23 standard deviations higher on the Educational Testing Service (ETS) test than did students in the comparison group. The difference is statistically significant. The average intervention group student performed at a level equivalent to that of a student in the 59th percentile of the comparison group.

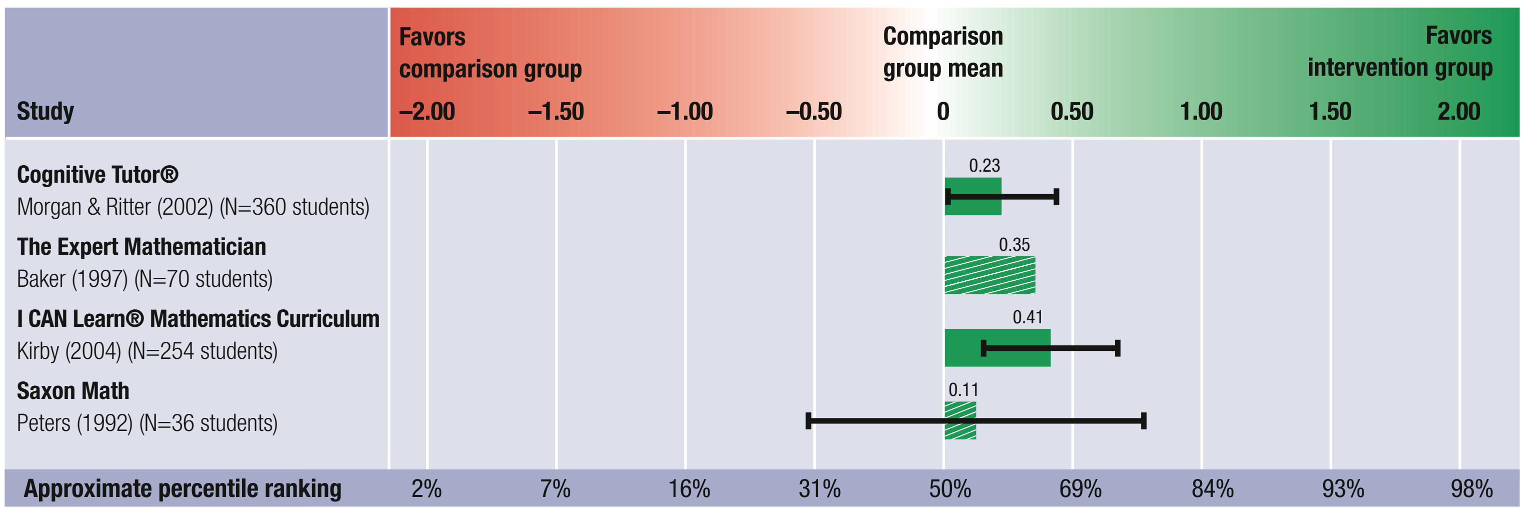
I CAN Learn® Mathematics Curriculum raised mathematics achievement of regular education students in 8th grade 0.41 standard deviations. This means that, on average, students in the intervention group scored at a level equivalent to that of students in the 66th percentile of the comparison group.

The RCT of *The Expert Mathematician* shows that it increased mathematics achievement compared to students in *Transition Mathematics;* however, it was not possible to compute statistical significance of this difference. The RCT of *Saxon Math* shows no evidence that the interventions enhanced mathematics achievement beyond that of students in the comparison group. The mean achievement test scores are not appreciably different between the intervention and comparison groups. The modest differences are no larger than could be expected from chance, given the variance in the outcome measures and the sample size (36 students).

Table 3

Estimated effects[a,b] of middle school math curricula (standard scores)

Randomized controlled trials without notable design or implementation flaws



a The topic report focuses on a limited number of math test scores per study. (All studies included at least one math test among their outcome measures.) In cases where a study reported on multiple outcomes, outcomes based on nationally normed tests were selected for inclusion over other measures; those based on state normed tests were selected over locally normed tests, unnormed tests, or other outcomes. Otherwise, unnormed tests were used. Where a study reported outcomes for students in successive school years, the topic report included those results for the most recent school year. Where the study reported outcomes for subgroups of students as well as for the overall sample, only results based on the overall sample were included in the topic report.

b Assumes a normal distribution.

How to read this table: The wide, shaded bar indicates both the direction and estimated size of the effect of the intervention. The estimated effects reported here are standardized differences in the mean values between the intervention and comparison groups. Bars extending to the right of zero denote estimated effects that favor the intervention group and those extending to the left of zero denote estimated effects that favor the comparison group. The solid line through the shaded bar marks the 95% confidence interval of the estimated effect; when there is no solid line, the study did not provide data to correctly compute the confidence interval. When the line does not cross zero (and the bar is solid, not striped), the estimate is statistically significant. The bar is striped if the effect is not significant or if the significance could not be accurately computed. The scale at the bottom of the chart indicates the approximate percentile distribution of students in the control group. The percentile ranking at the end of the shaded bar can be used to interpret the standardized mean difference in the outcome. For example, an effect of .5 is roughly equivalent to an increase in the mean value from that of the average student in the comparison group (50th percentile) to that of the average student at the 69th percentile.

Estimated effects: randomized controlled trials *(continued)*

Randomized controlled trials

Cognitive Tutor®

The RCT of *Cognitive Tutor®* compared it with *McDougal Littell's Health Algebra I* curriculum directed toward 9th-grade regular education students in five schools in Oklahoma. The main outcome variables included a standardized test, the Education Testing Service (ETS) End of Course Algebra Assessment, and first and second semester grades.

I CAN Learn® Mathematics Curriculum

The RCT of the *I CAN Learn® Mathematics Curriculum* involved 8th-graders in a middle school in Georgia's Gilmer County School District. Students were randomly assigned to the *I CAN Learn® Mathematics Curriculum* or to traditional instruction.

The Expert Mathematician

The RCT of *The Expert Mathematician* involved 8th-grade students randomly assigned to the intervention group or to *Transition Mathematics*, part of the *University of Chicago School Mathematics Project*. The study was conducted in a St. Louis, Missouri, middle school serving largely students from low-income families.

Saxon Math

The RCT of *Saxon Math* involved math-talented 8th-grade students randomly assigned to *Saxon Math* or to the National Council of Teachers of Mathematics (NCTM) standards-based *University of Chicago School Mathematics Project* curriculum. Students in the study sample attended a junior high school in Lincoln, Nebraska.

Estimated effects: quasi-experimental design studies

Two interventions were evaluated using both RCTs and QEDs. For one, *I CAN Learn® Mathematics Curriculum*, results of the QED studies are consistent with those for the RCT. The results for the other, *Saxon Math*, differ (table 4).

The two QED studies on *I CAN Learn® Mathematics Curriculum* showed evidence that the intervention improved mathematics achievement above what students would have attained otherwise. However, the findings were not statistically significant for one study (Kerstyn 2001) and significance could not be accurately computed for the other (Brooks 1999).

The QED study of *Saxon Math* (Crawford and Raia 1986), reports an increase (0.41 standard deviations) in mathematics achievement. This means that the average student in *Saxon Math* performed at the same level as students at the 66th percentile in the comparison group. However, the significance of this difference cannot be accurately computed. In comparison, the one RCT for this intervention (Peters 1992) indicated that mathematics achievement was not significantly higher among students using *Saxon Math* than it would have been had they remained in the regular school mathematics curriculum.

The estimated effects of the *Connected Mathematics Project*, based on QED studies, differ from one study report to the next. Estimates from Ridgeway (2002) based on the Iowa Test of Basic Skills (ITBS) are positive for students in 8th grade. But, the study does not provide sufficient information to determine whether the estimated impact is statistically significant. The average estimated effects are not significantly different from zero in either the Riordan and Noyce (2001) study using the Massachusetts Comprehensive Assessment System (MCAS) exam (grades 6–8) or the Schneider (2000) study of a sample of Texas schools serving students in grades 6–8, using the Texas Assessment of Academic Skills (TAAS).

Quasi-experimental design studies

Connected Mathematics Project

The three QED studies involved 6th- and 7th-graders in three regions of the country, 8th graders in Massachusetts schools, and schools in rural, suburban, and urban areas of Texas.

I CAN Learn® Mathematics Curriculum

The QED studies of *I CAN Learn® Mathematics Curriculum* involved Title I middle schools in Florida and 7th- through 10th-grade students in five Louisiana districts.

Saxon Math

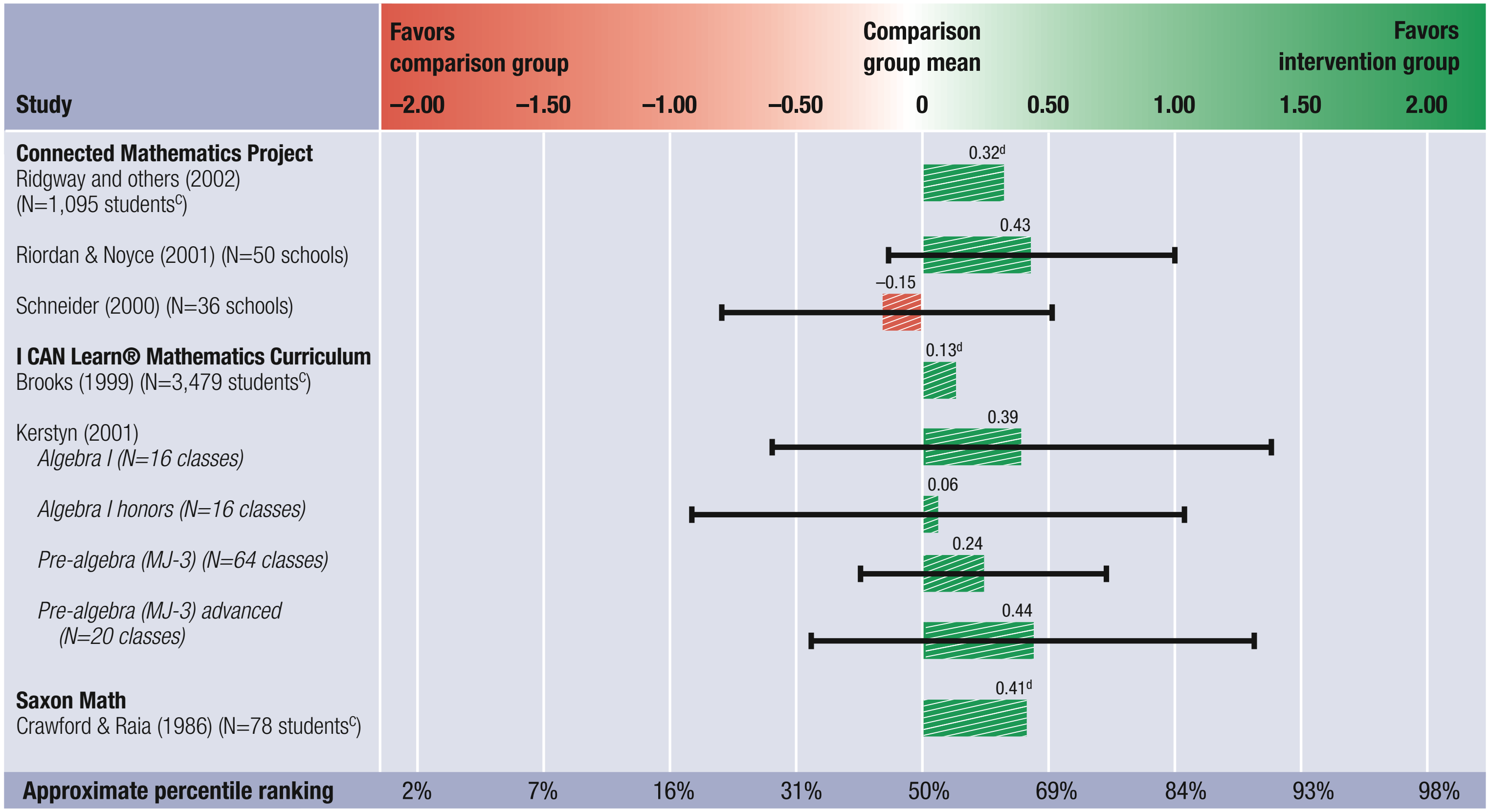
The one QED study involved 8th-grade students in Oklahoma City middle schools.

Estimated effects: quasi-experimental design studies *(continued)*

Table 4

Estimated effects[a,b] of middle school math curricula (standard scores and percentiles)

Quasi-experimental design studies and/or randomized controlled trials with notable flaws



a The topic report focuses on a limited number of math test scores per study. (All studies included at least one math test among their outcome measures.) In cases where a study reported on multiple outcomes, outcomes based on nationally normed tests were selected for inclusion over other measures; those based on state normed tests were selected over locally normed tests, unnormed tests, or other outcomes. Otherwise, unnormed tests were used. Where a study reported outcomes for students in successive school years, the topic report included those results for the most recent school year. Where the study reported outcomes for subgroups of students as well as for the overall sample, only results based on the overall sample were included in the topic report.
b Assumes a normal distribution.
c Sample size reported is unit of analysis, not unit of assignment.
d When there is no solid line, the study did not provide data to correctly compute the confidence interval.
How to read this table: The wide, shaded bar indicates both the direction and estimated size of the effect of the intervention. The estimated effects reported here are standardized differences in the mean values between the intervention and comparison groups. Bars extending to the right of zero denote estimated effects that favor the intervention group and those extending to the left of zero denote estimated effects that favor the comparison group. The solid line through the shaded bar marks the 95% confidence interval of the estimated effect. When the line does not cross zero (and the bar is solid, not striped), the estimate is statistically significant. The bar is striped if the effect is not significant or if significance could not be accurately computed. The scale at the bottom of the chart indicates the approximate percentile distribution of students in the control group. The percentile ranking at the end of the shaded bar can be used to interpret the standardized mean difference in the outcome. For example, an effect of 0.5 is roughly equivalent to an increase in the mean value from that of the average student in the comparison group (50th percentile) to that of the average student at the 69th percentile.

Call for new evidence

The evidence base is sparse. There have been few RCTs—the best evidence possible for causal inferences—and those few typically are small. The RCTs can be improved by testing the intervention in multiple sites to learn whether it works in different places. Both randomized experiments and quasi-experiments can be improved by including key subgroups in the study's design and analysis—such as gender, race/ethnicity, and economic status. Most of the studies could be improved through more uniform and complete reporting—especially on the implementation of interventions and on the instruction to comparison group students—and through more adequate statistical analysis and reporting.

References

Baker, J. J. (1997). *Effects of a generative instructional design strategy on learning mathematics and on attitudes towards achievement*. Unpublished doctoral dissertation, University of Minnesota, Minneapolis. *Dissertations Abstracts International, 58* (7), 2573A. (UMI No. 980095)

Kirby, P. C. (2004). *Comparison of I CAN Learn and traditionally-taught 8th grade student performance on the Georgia Criterion-Referenced Competency Test*. Unpublished manuscript.

Morgan, P., & Ritter, S. (2002). *An experimental study of the effects of Cognitive Tutor® Algebra I on student knowledge and attitude*. (Available from Carnegie Learning, Inc., 1200 Penn Avenue, Suite 150, Pittsburgh, PA 15222)

Peters, K. G. (1992). *Skill performance comparability of two algebra programs on an eighth-grade population.* Unpublished doctoral dissertation, University of Nebraska, Lincoln, NE.

Brooks, C. (1999, August). *Evaluation of Jefferson Parish Technology Grant I CAN Learn® Algebra I*. (Available from the Department of Educational Leadership, University of New Orleans, New Orleans, LA 70148)

Crawford, J., & Raia, F. (1986, February). *Analyses of eighth grade math texts and achievement (evaluation report).* Oklahoma City: Planning, Research, and Evaluation Department, Oklahoma City Public Schools.

Hoover, M., Zawojewski, J. D., & Ridgway, J. E. (1997, April). *Effects of the Connected Mathematics Project on student attainment*. Paper presented at the annual meeting of the American Education Research Association, San Diego, CA.]

Kerstyn, C. (2001). *Evaluation of the I CAN Learn® mathematics classroom: First year of implementation (2000–2001 school year).* (Available from the Division of Instruction, Hillsborough County Public Schools, Tampa, FL)

Ridgway, J. E., Zawojewski, J. S., Hoover, M. N., & Lambdin, D. V. (2002). Student attainment in Connected Mathematics curriculum. In S. L. Senk & D. R. Thompson (Eds.), *Standards-based school mathematics curricula: What they are? What do students learn?* Mahwah, NJ: Erlbaum.

Riordan, J. & Noyce, P. (2001). The impact of two standards-based mathematics curricula on student achievement in Massachusetts. *Journal for Research in Mathematics Education, 32* (4), 368–398.

Schneider, C. L. (2000). *Connected Mathematics and the Texas Assessment of Academic Skills*. Unpublished doctoral dissertation, University of Texas, Austin, TX.

National Commission on Excellence in Education. (1983). *A nation at risk: The imperative for educational reform.* Washington, DC: U.S. Government Printing Office.

National Research Council. (1990). *A challenge of numbers: People in the mathematical sciences*. Washington, DC: National Academy Press.

The What Works Clearinghouse (www.whatworks.ed.gov) was established in 2002 by the U.S. Department of Education's Institute of Education Sciences to provide educators, policymakers, researchers, and the public with a central and trusted source of scientific evidence of what works in education. Please email all questions and comments to info@whatworks.ed.gov. The What Works Clearinghouse is administered by the U.S. Department of Education through a contract to a joint venture of the American Institutes for Research and the Campbell Collaboration.

Appendix

Table A1 Middle school math curricula considered for this review

Curricula for which there is at least one study that meets WWC evidence standards	Curricula for which no studies were found
Cognitive Tutor®	Addison-Wesley Mathematics basal program
Connected Mathematics Project	AGS Publishing
The Expert Mathematician	A+ny where Learning System
I CAN Learn® Mathematics Curriculum	Academic Systems' Interactive Math Curriculum
Saxon Math	Math Applications and Connections (textbook published by Glencoe)
Curricula for which all identified studies failed to meet WWC evidence standards	Heath Mathematics Connections (textbook series)
	Heath Math Passport (textbook series)
Algebra Project	Holt Middle School Math (textbook)
Algebraic Thinking	Key Math Teach and Practice
Compass Learning	Larson Developmental Math Series
Connecting Math Concepts (CMC)	Lightspan Achieve Now
CORD Applied Math	Math Advantage (textbook series)
Destination Math	Math Learning Center
Integrated Mathematics Science and Technology (IMaST)	Mathematics Plus (textbook series published by Harcourt)
Jasper Series	MathScape: Seeing and Thinking Mathematically
Mathematics in Context	Macmilllan/McGraw-Hill
MATHThematics	Middle Grades Math (textbook series, published by ScottForesman/AddisonWesley)
Model Mathematics Program (MMP)	Middle School Math (textbook series)
Moving with Math	Middle School Mathematics through Applications Program (MMAP)
Partnerships for Access to Higher Mathematics (PATH Mathematics)	The New Century Integrated Instructional System
PLATO	Real Math basal mathematics program
Successmaker	Scott Foresman Math Diagnostic & Intervention System
	SimCalc: Cognitive Foundations for a Multiplicative Structures Curriculum
	Singapore Mathematics

Notes: Two studies look at home-grown interventions, which are not in this list of published, commercial interventions.
The University of Chicago School Mathematics Project is currently under review.

Table A2 Description of studies of middle school curricula

Evidence base rating	Characteristic	Description
Cognitive Tutor®		
Morgan, P., & Ritter, S. (2002). *An experimental study of the effects of Cognitive Tutor Algebra I on student knowledge and attitude.* (Available from Carnegie Learning, Inc., 1200 Penn Avenue, Suite 150, Pittsburgh, PA 15222)		
✓✓	Participants	369 9th-grade students in the part of the study that used random assignment (with 360 students measured on the ETS test). Approximately two-thirds of the students in the study were white, and the other one-third was composed of students who were Asian, Black, Hispanic, Native American, and other. None of the students was in special education.The WWC review focuses on the subsample that was randomly assigned to conditions, using additional information provided by the study author.
	Setting	Four urban junior high schools in the Moore Independent School District in Oklahoma (in the part of the study that used random assignment). Six teachers in three of four intervention schools taught both the intervention and the comparison curricula.
	Intervention	Students in the intervention group were taught using *Cognitive Tutor®* for one academic year. Students spent 60 percent of each class period in group activities and classroom discussions and their remaining time working with *Cognitive Tutor®* to devleop their own problem-solving skills.
	Comparison	Students in the comparison group were taught using the *McDougal Littell's Heath Algebra I,* which is a traditional, teacher-directed curriculum.
	Primary outcomes and measurement	ETS Algebra I End-of-Course Assessment, developed by ETS, consisting of 25 multiple-choice and 15 constructed-response questions. The other two outcomes were semester 1 grades and semester 2 grades.
	Teacher training	Teachers who taught *Cognitive Tutor®* took a standard four-day training right before the intervention began.
Connected Mathematics Project		
Ridgway, J.E., Zawojewski, J.S., Hoover, M.N, & Lambdin, D.V. (2002). Student attainment in connected mathematics curriculum. In S.L. Senk & D R. Thompson (Eds.), *Standards-based school mathematics curricula: What they are? What do students learn?* Mahwah, NJ: Erlbaum.		
✓	Participants	The 1994/95 sample included 338 6th-graders and 627 7th-graders from nine *Connected Mathematics Project* schools (two classrooms per grade from each school), and 162 6th-graders and 234 7th-graders from nine comparison schools (one classroom per grade from each school). The 1995/96 sample included 820 8th-graders from an unspecified number of *Connected Mathematics Project* schools and 275 8th-graders from an unspecified number of comparison schools. Some students were included in both the 1994/95 sample and the 1995/96 sample. Demographic characteristics of the participants are not reported.
	Setting	Participating classrooms were from schools located in the Midwest, West, and East regions of the country.
	Intervention	Teachers in the intervention group were using *Connected Mathematics Project* as the core curriculum throughout the school year. The study authors do not report, however, how *Connected Mathematics Project* was actually implemented in those classrooms. Student participants received varying amount of intervention. All the 6th-grade students in the study were new to *Connected Mathematics Project*, and about three-fourths of the 7th-grade and 8th-grade students in the study had used *Connected Mathematics Project* in the previous year.

Table A2 **Description of studies of middle school curricula** *(continued)*

Evidence base rating	Characteristic	Description
	Comparison	Teachers in the comparison group did not implement *Connected Mathematics Project*, nor were they involved in any reform efforts. Data were not available about the mathematics textbook series used by those teachers.
	Primary outcomes and measurement	Iowa Test of Basic Skills (ITBS) Survey Battery and Balanced Assessment (BA) Test. ITBS is a norm-referenced standardized test. BA is a test designed to assess students' math achievement in a variety of curricular areas through constructed-response items. It was developed through the collaboration between the *Connected Mathematics Project* developer and the Balanced Assessment Project.
	Teacher training	All *Connected Mathematics Project* teachers attended the summer Connected Mathematics Project institutes at Michigan State University.
Riordan, J. & Noyce, P. (2001). The impact of two standards-based mathematics curricula on student achievement in Massachusetts. *Journal for Research in Mathematics Education, 32* (4), 368–398.		
✓	Participants	20 *Connected Mathematics Project* schools with 1,879 8th-graders and 30 matched comparison schools with 4,978 8th-graders. Overall, 10 percent of the student participants were eligible for free or reduced-price lunch, and 87 percent of the students were white. All students were regular education students.
	Setting	Relatively advantaged middle schools with predominantly white students and a low percentage of students receiving free or reduced-price lunch in the state of Massachusetts
	Intervention	Schools in the intervention group had implemented at least 11 student units in grades 6–8 by 1998/99, but none of the schools implemented all eight units that *Connected Mathematics Project* has available for each grade. Further, it is not clear how *Connected Mathematics Project* was actually implemented in those schools. Twenty schools in the intervention group had implemented *Connected Mathematics Project* for two to three years, and one school had implemented the program for four years.
	Comparison	The 30 comparison schools did not implement *Connected Mathematics Project* but used 15 different textbook programs, which, in the aggregate, represented the instructional norm in Massachusetts. The most commonly used programs were those published by Heath, Addison-Wesley, Prentice Hall, and Houghton-Mifflin.
	Primary outcomes and measurement	Massachusetts Comprehensive Assessment System, a criterion-referenced state test that includes both multiple-choice and open-response questions.
	Teacher training	No teacher training reported.
Schneider, C.L. (2000). Connected Mathematics and the Texas Assessment of Academic Skills. *Dissertation Abstracts International, 62* (02), 503. (UMI No. 3004374)		
✓	Participants	23 *Connected Mathematics Project* schools and 25 matched comparison schools, overall, including three smaller cohorts. Cohort 1 had 23 intervention and 19 comparison schools. Cohort 2 had 22 intervention and 19 comparison schools. Cohort 3 had 18 intervention and 18 comparison schools. Those schools varied in the racial composition, socioeconomic status, special education status, and English language learner status of the student populations that they served. Many of the schools had predominantly minority student populations.

Table A2 **Description of studies of middle school curricula** *(continued)*

Evidence base rating	Characteristic	Description
	Setting	Schools in rural, suburban, and urban, and both low and high socioeconomic areas of Texas.
	Intervention	Schools in the intervention group were using *Connected Mathematics Project* for grades 6–8. There were substantial variations in the extent to which the curriculum was used at each grade and each year across these schools. The three cohorts in the intervention group received *Connected Mathematics Project* for three years, two years, and one year respectively between 1996/97 and 1998/99.
	Comparison	The comparison schools did not implement *Connected Mathematics Project*, and it is unclear what mathematics curricula they were using.
	Primary outcomes and measurement	Texas Assessment of Academic Skills (TAAS) passing rate and Texas Learning Index (TLI). TAAS is a criterion-referenced state test that measures problem-solving and critical-thinking skills. TLI is a TAAS-based statistic designed for comparing student progress between administrations and between grades.
	Teacher training	Teachers who taught grade 6, 7, or 8 at the 23 *Connected Mathematics Project* schools participated in a six-day summer professional development conducted by the Texas Statewide Systemic Initiative in 1996, 1997, and 1998.
The Expert Mathematician **Baker J.J. (1997). *Effects of a generative instructional design strategy on learning mathematics and on attitudes towards achievement.* Unpublished doctoral dissertation, University of Minnesota. *Dissertations Abstracts International, 58* (7), 2573A. (UMI No. 9800955)**		
✓✓	Participants	70 8th-grade students. Most were from low-income families and qualified for free or reduced-price lunch. All but three students were white. None was in special education. Students were randomized to the intervention or comparison condition.
	Setting	Surburban middle school in St. Louis, Missouri; two intervention classrooms and two comparison classrooms.
	Intervention	The intervention group experienced a generative mathematics curriculum that used *The Expert Mathematician* (version 3.0). Students worked individually or in pairs with the printed materials and the computer to work through the curriculum lessons. Lessons were 85 minutes long, every other day for one school year. The same teacher taught both the intervention and the comparison students.
	Comparison	The comparison group experienced a linear mathematics curriculum based on *Transition Mathematics*, a traditional, teacher-directed curriculum.
	Primary outcomes and measurement	78-item Objectives by Strand math test developed by the school district and administered at the end of the school year by the classroom teacher. No norming information is available.
I CAN Learn® Mathematics Curriculum **Kirby, P. C. (2004). *Comparison of I CAN Learn® and traditionally-taught 8th grade student performance on the Georgia Criterion-Referenced Competency Test.* Unpublished manuscript.**		
✓✓	Participants	254 8th-grade regular education students, 91 of whom were randomly assigned to the intervention group and 163 of whom were in the comparison group.

Evidence base rating	Characteristic	Description
	Setting	One middle school in the Gilmer County school district in northwestern Georgia.
	Intervention	Students in intervention group had the *I CAN Learn® Mathematics Curriculum* for one school-year. This is a software-based curriculum that follows a five-part format consisting of pretest, review, lesson presentation, quiz, and cumulative review. The software provides feedback, homework, and assessment. There are a total of 109 algebra lessons. All students in the intervention group were taught by the same teacher.
	Comparison	Students in the comparison group had a traditional math curriculum, the nature of which is not specified in the study report.
	Primary outcomes and measurement	Georgia Criterion-Referenced Competency Test (GCRCT), a state test. The math test contains 60 items in six areas: number sense and numeration, geometry and measurement, patterns and relationships/algebra, statistics and probability, computation and estimation, and problem solving.
	Teacher training	There was only one *I CAN Learn®* teacher in this study. She was trained the prior year.
Kerstyn, C. (2001). ***Evaluation of the I CAN Learn® mathematics classroom: First year of implementation (2000–2001 school year).*** **(Available from the Division of Instruction, Hillsborough County Public Schools, Tampa, FL)**		
✓	Participants	8th-grade students in Title I middle schools. The study was limited to regular education students. The students were racially diverse and many were eligible for free or reduced-priced lunch. There were 58 intervention classrooms and 58 comparison classes used in the analysis.
	Setting	36 middle schools in the Hillsborough County Public School system in Florida. This county includes the Tampa metropolitan area.
	Intervention	The *I CAN Learn® Mathematics Curriculum*. Each *I CAN Learn®* lesson follows a five-part format consisting of pretest, review, lesson presentation, quiz, and cumulative review. The *I CAN Learn®* program is software-based and provides feedback, homework, and assessment. The *I CAN Learn®* curriculum consists of 109 algebra lessons but the author does not indicate how many of them are required to be completed in order for the curriculum to be implemented as intended. The author indicates that the *I CAN Learn®* curriculum was implemented in class periods of 45, 50, 80, and 90 minutes in length. When surveyed, the teachers reported that 45 minutes was not long enough to make it through the curriculum. In this study, the intervention lasted for one school year.
	Comparison	Classes using a traditional math curriculum. The author does not provide further information on the curriculum.
	Primary outcomes and measurement	Florida Comprehensive Assessment Test (FCAT), a state test and MJ-3 Cumulative Test (an end of semester exam administered at the end of the first semester). The MJ-3 Cumulative Test is not a state or nationally normed, standardized test but it has established reliability.
	Teacher training	Teachers in this study participated in training sessions on the use of the software and hardware, but not on use of the software in instruction.

Table A2 **Description of studies of middle school curricula** *(continued)*

Evidence base rating	Characteristic	Description
Brooks, C. (1999). *Evaluation of Jefferson Parish Technology grant: I CAN Learn® Algebra I.* Unpublished reported submitted to the Superintendent of Jefferson Parish Public Schools. (Available from the Department of Education Leadership, University of New Orleans, New Orleans, LA 70148)		
✓	Participants	3,479 students enrolled in 7th through 10th grade and came from honors, gifted, and remedial classes. There were 102 intervention classes and 67 comparison classes at the start of the study.
	Setting	21 schools (20 parochial, 1 public) with five school districts in Louisiana.
	Intervention	The *I CAN Learn® Mathematics Curriculum*. Each *I CAN Learn®* lesson follows a five-part format consisting of pretest, review, lesson presentation, quiz, and cumulative review. The *I CAN Learn®* program is software-based and provides feedback, homework, and assessment. The *I CAN Learn®* curriculum consists of 109 algebra lessons, but the author does not indicate how many of them are required to be completed in order for the curriculum to be implemented as intended. In this study, the intervention lasted for one school year.
	Comparison	Traditionally taught. No other information provided.
	Primary outcomes and measurement	30-item criterion-referenced achievement test developed by the evaluation team based on student textbooks. This test is not a state or nationally normed, standardized test. However, it has established reliability.
	Teacher training	Teachers received training before the start of the 1998/99 school year.
Saxon Math		
Peters, K. G. (1992). *Skill performance comparability of two algebra programs on an eighth-grade population.* Unpublished doctoral dissertation, University of Nebraska, Lincoln, NE.		
✓✓	Participants	36 8th-grade students. All of the students were math talented based on teacher recommendations, prior academic achievement, and personal maturity level.
	Setting	One junior high school in a rural-suburban district near Lincoln, Nebraska.
	Intervention	Students in the intervention group were taught using the *Saxon Math* curriculum for 8th-grade students (Algebra 1/2). They participated in 60-minute daily sessions throughout one school year. Each session began with the teacher introducing a new concept. Students then had the opportunity to practice both the new concept and concepts that had been introduced in previous sessions. Students were assessed every fifth lesson. The intervention is designed to cover 120 lessons across a one-year period. Students participate in daily lessons, approximately 60 minutes per lesson.
	Comparison	Students in the comparison group were taught using an NCTM standards based curriculum called the *University of Chicago Mathematics Project*. This curriculum was designed to build independent learners and thinkers and to develop understanding of math vocabulary (such as mathematical signs). It emphasizes review of concepts within existing lessons as a means to increase student comprehension. Comparison group students were taught by the same teachers as were those in the intervention group.

Table A2 **Description of studies of middle school curricula** *(continued)*

Evidence base rating	Characteristic	Description
	Primary outcomes and measurement	The primary outcome measure is the Orleans-Hannah Algebra Prognosis Test, a nationally normed 60-item test designed to predict student success in future algebra study.
	Teacher training	No special training of teachers in the use of *Saxon Math* was discussed. A number of teacher resources are available via the *Saxon* website, including telephone and email access to customer service and educational representatives (in each state).
Crawford, J., & Raia, F. (1986, February). ***Analyses of eighth grade math texts and achievement (evaluation report).*** **Oklahoma City: Planning, Research, and Evaluation Department, Oklahoma City Public Schools.**		
✓	Participants	78 8th-grade students matched on pre-test California Achievement Test (CAT) scores.
	Setting	Four middle schools in the Oklahoma City Public Schools; four teachers who taught both the intervention and the comparison curricula.
	Intervention	Participants in the intervention group were taught using the *Saxon Math* curriculum for 8th-grade students (Algebra 1/2). Specific information about the level of implementation was not provided by Crawford and Raia (1986). The intervention is designed to cover 120 lessons across a one-year period with students participating in daily lessons, approximately 60 minutes per lesson. Students participated during the 1984/85 academic year.
	Comparison	Participants in the comparison group were taught using the Scott-Foresman Mathematics curriculum. Information about this curriculum, including implementation, was not provided by Crawford and Raia (1986).
	Primary outcomes and measurement	The primary outcome measure is the California Achievement Test (CAT), including overall scores and scores for math concepts and math computation. The CAT is a nationally normed, valid and reliable test designed to measure achievement in the basic skills taught in school.
	Teacher training	Teacher training was not reported for this study; however, a number of teacher resources are available via the *Saxon* website, including telephone and email access to customer service and educational representatives (in each state).

Table A3 **Estimated effects of curriculum-based interventions designed to increase math achievement among middle school students: results based on well designed and implemented randomized controlled trials**

				Mean outcome		*Standard deviation*[a]		*Estimated impact*[b]	
Study	***Study sample***	***Measure***	***Sample size***	***Intervention group***	***Comparison group***	***Intervention group***	***Comparison group***	***Mean difference***	***Standardized difference in means***
Cognitive Tutor®									
Morgan and Ritter (2002)	9th-grade regular education students	Educational Testing Service (ETS) Algebra	360 students	16.7	15.4	5.70	5.60	1.3	0.23 (±0.21)**
The Expert Mathematician									
Baker (1997)[c]	8th-grade regular education students	Objectives by Strand test	70 students	45.1	40.8	12.03	12.41	4.3	0.35
I CAN Learn® Mathematics Curriculum									
Kirby (2004)	8th-grade regular education students	GCRCT[d]	254 students	333.5	319.9	35.70	31.70	13.6	0.41 (±0.26)***
Saxon Math									
Peters (1992)	8th-grade "math-talented" students	Achievement Test	36 students	95.6	95.1	4.53	4.09	0.5	0.11 (±0.65)

****** Statistically significant at the 1 percent level, two-tailed test. ******* Statistically significant at the 5 percent level, two-tailed test.

a Shows how dispersed the participants' outcomes are. A small standard deviation would suggest that participants had similar outcomes.

b The WWC estimated the impact based on statistics reported by the study author.

c Means and effects adjusted for pretest differences between the groups.

d Georgia Criterion-Referenced Competency Test.

Table A4 **Estimated effects of curriculum-based interventions designed to increase math achievement among middle school students: results based on quasi-experimental design studies and randomized controlled trials with design or implementation flaws**

Study	Study sample	Measure	Sample size	Mean outcome: Intervention group	Mean outcome: Comparison group	Standard deviation[a]: Intervention group	Standard deviation[a]: Comparison group	Estimated impact[b]: Mean difference	Estimated impact[b]: Standardized difference in means
Connected Mathematics Project									
✓ Ridgway and others (2002)[c]	8th grade; midwestern middle schools	Iowa Test of Basic Skills (ITBS)	1,095 students	9.4	8.6	2.5	2.6	0.8	0.32
✓ Riordan & Noyce (2001)	8th grade largely nonpoor	Comprehensive Assessment System (MCAS)	50 schools	238.2	233.9	9.1	10.3	4.3	0.43 (±1.51)
✓ Schneider (2000)	6th–8th grade varied background, many minorities	Texas Assessment of Academic Skills (TAAS)	36 schools	73.6	74.5	5.41	6.37	–0.9	–0.15 (±0.65)
I CAN Learn® Mathematics Curriculum									
✓ Brooks (1999)[c]	7th- through 10th-grade honors, gifted, and remedial students	30-item criterion reference algebra test)	3,479 students	7.5	6.9	4.7	4.0	0.6	0.13
✓ Kerstyn (2001)	Algebra 1 classes, Title I middle school	Florida Comprehensive Assessment Test (FCAT)	16 classes	351.1	345.4	15.6	11.4	5.7	0.39 (±0.99)
	Algebra 1 honors, Title I middle school	FCAT	16 classes	374.2	373.1	11.1	20.5	1.1	0.06 (±0.98)
	Pre-algebra (MJ-3), Title I middle school	FCAT	64 classes	298.0	294.4	15.6	13.8	5.4	0.24 (±0.49)
	Pre-algebra (MJ-3) advanced, Title I middle school	FCAT	20 classes	331.5	326.1	12.6	11.0	5.4	0.44 (±0.89)
Saxon Math									
✓ Crawford & Raia (1986)[c,d]	8th grade, urban middle school	California Achievement Test (CAT) Overall Math	78 students	55.6	50.7	11.9	11.8	4.8	0.41

****** Statistically significant at the 1 percent level, two-tailed test.
a Shows how dispersed the participants' outcomes are. A small standard deviation would suggest that participants had similar outcomes.
b The WWC estimated the impact based on statistics reported by the study author.
c The statistical significance of the impact estimate reported by the study author was improperly calculated. It ignored the fact that classrooms or schools, not students, were the unit of random assignment. The study did not provide sufficient data to compute the correct standard error of the impact estimate. The sample size reported is in unit of analysis, not unit of assignment.
d Means and effects adjusted for pretest differences between groups.